Supplemental materials for

Bryan CJ, Allen MH, Thomsen CJ, et al. Improving suicide risk screening to identify the highest risk patients: results from the PRImary Care Screening Methods (PRISM) study. *Ann Fam Med*. 2021;19(6):492-498.

Supplemental Appendix
Information about Missing Data and Multiple Imputation Methods

Follow-up assessment data were missing from 952 of 2744 (34.7%) participants. Each participant was contacted 6 and 12 months after enrollment up to three times by phone, three times by email, and two times by mail to schedule follow-up assessments. Participants who did not respond to these multiple contact attempts at both assessments were considered lost to follow-up (n=765, 27.9% of the sample). Another 186 participants (6.8% of sample) withdrew from the study. One participant (0.04% of the sample) died of natural causes in his sleep after a medical procedure unrelated to the study. We learned of no suicide deaths by participants during the course of the study.

Prior to our analyses, we evaluated missing data patterns and found that participants with missing follow-up data were younger, less likely to be white, and had significantly higher SCS scores at baseline:

Variable	Not Missing	Missing	Test
Modified SCS total score	6.2	7.2	t(1637)=2.1, p=.035
PHQ item 9	0.1	0.2	t(1729)=1.1, p=.291
PHQ-8 total score	6.0	5.9	t(1800)=0.4, p=.657
Age	42.2	37.0	t(1946)=6.7, p<.001
Prior suicide attempt	9.2	9.0	$\chi^2(2)=0.0$, p=.888
Gender (male)	53.2	49.3	$\chi^2(2)=3.6$, p=.058
White race	70.2	62.5	$\chi^2(2)=16.4$, p<.001

Rubin[1] described three types of missing data that differ with respect how the probability of a variable's missingness is related (or not) to other measured variables. Data are considered missing completely at random (MCAR) when a variable's missingness is not correlated with other observed variables and is unrelated to the variable itself. Data are considered missing at random (MAR) when a variable's missingness is correlated with other observed variables but is not related to the variable itself. Data are considered missing not at random (MNAR) when a variable's missingness is related to the variable itself.

The results of our missing data analysis indicated that missing follow-up data (i.e., suicidal behaviors) were significantly correlated with other observed variables used in our study—age, race, and SCS total score—indicating that MCAR is not supported. The observed rate of suicidal behavior during follow-up among participants with complete data (0.5% at 1 month and 1.0% at 3 months) were consistent with the a priori estimates used for our power analysis, suggesting missingness was unlikely to be related to the occurrence of suicidal behavior itself and suggesting that MNAR did not characterize the data. We therefore approached our analyses on the assumption of MAR, a condition well-suited for multiple imputation.

We imputed 10 datasets using all observed variables from the baseline assessment and then pooled parameter estimates to derive regression coefficients, standard errors, confidence intervals, and p-values, consistent with the recommendations and procedures described by Rubin.[2] We next repeated all analyses using only complete cases (i.e., participants without missing data) and compared the results of those models to the pooled results. The results derived from the complete dataset did not differ meaningfully from those derived from the pooled imputed datasets. We therefore reported the results of the pooled analyses in the text.

Supplemental Tables

Supplemental Table 1.

Frequency of PHQ-9 item 9 responses at baseline among primary care patients with suicidal behavior during the first 30 and days 90 postbaseline

	1 Month (n=13)			3 Months (n=28)	
PHQ-9 Item 9 Response	n	(%)		n	(%)
Not At All	5	(38.5)		11	(39.3)
Several Days	4	(30.1)		8	(28.6)
More than half the days	1	(7.7)		3	(10.7)
Nearly every day	3	(23.1)		5	(17.9)
Missing	0	(0.0)		1	(3.6)

Supplemental Table 2.

Pooled intercorrelations among predictor variables across the 10 imputed datasets

Variable	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**
1. Modified SCS total score	--						
2. PHQ item 9	0.58***	--					
3. Prior suicide attempt	0.34***	0.28***	--				
4. PHQ-8 total score	0.63***	0.54***	0.32***	--			
5. Age	-0.02	-0.02	-0.03	-0.02	--		
6. Gender (male)	0.06**	0.01	0.09***	0.09***	-0.01	--	
7. White race	-0.01	0.00	0.03	-0.04	-0.01	-0.11***	--

Supplemental Table 3.

Pooled results of univariate and multivariate logistic regression models predicting suicidal behaviors within 30 and 90 days postbaseline

		Univariate				Multivariate[b]	
	OR	(95% CI)	p		aOR	(95% CI)	p
30 days							
Modified SCS total score	1.10	(1.04, 1.16)	.005		1.06	(0.98, 1.16)	.169
PHQ item 9	3.26	(1.57, 6.76)	.005		0.80	(0.20, 3.12)	.748
Prior suicide attempt	16.46	(2.07, 131.16)	.013		6.38	(0.52, 78.01)	.157
PHQ-8 total score	1.21	(1.07, 1.37)	.005		1.11	(0.94, 1.31)	.219
Age	1.01	(0.96, 1.06)	.750		1.04	(0.96, 1.12)	.310
Gender	0.73	(0.12, 4.48)	.736		0.71	(0.06, 8.16)	.781
White race [a]	--	--	--		--	--	--
90 days							
Modified SCS total score	1.10	(1.07, 1.13)	<.001		1.08	(1.03, 1.13)	.002
PHQ item 9	3.34	(2.24, 4.98)	<.001		1.02	(0.48, 2.15)	.963
Prior suicide attempt	7.80	(2.52, 24.14)	.001		1.92	(0.50, 7.45)	.347
PHQ-8 total score	1.19	(1.10, 1.28)	<.001		1.03	(0.91, 1.15)	.662
Age	0.99	(0.95, 1.02)	.531		1.00	(0.97, 1.04)	.861
Gender	1.17	(0.44, 3.10)	.747		1.46	(0.45, 4.74)	.528
White race	1.67	(0.52, 5.40)	.394		1.67	(0.51, 7.62)	.324

[a] Variable excluded from model due to low cell counts resulting in quasi-partial separation of the model.
[b] Across the 10 imputed datasets, the pooled Nagelkerke R^2 values were 0.30 for the first 30 days postbaseline and 0.26 for the first 90 days postbaseline.

Supplemental Table 4.

Area under the curve (AUC) values for PHQ-8 and SCS items as indicators of suicidal behavior during the first 30 and 90 days postbaseline among individuals screening positive for suicide risk at baseline

	30 days postbaseline			90 days postbaseline	
Scale Item	AUC	(95% CI)		AUC	(95% CI)
PHQ-8 Item					
1	0.577	(0.366, 0.787)		0.620	(0.479, 0.762)
2	0.626	(0.383, 0.869)		0.616	(0.463, 0.768)
3	0.678	(0.496, 0.861)		0.613	(0.476, 0.750)
4	0.629	(0.454, 0.803)		0.567	(0.421, 0.712)
5	0.693	(0.525, 0.860)		0.620	(0.473, 0.767)
6	0.562	(0.328, 0.796)		0.615	(0.465, 0.764)
7	0.590	(0.367, 0.813)		0.579	(0.422, 0.736)
8	0.598	(0.343, 0.852)		0.626	(0.472, 0.781)
SCS Item [a]					
1	0.730	(0.749, 0.964)		0.681	(0.712, 0.915)
3	0.674	(0.720, 0.932)		0.659	(0.686, 0.889)
4	0.604	(0.636, 0.947)		0.641	(0.680, 0.894)
5	0.741	(0.727, 0.952)		0.738	(0.717, 0.917)
6	0.644	(0.738, 0.973)		0.648	(0.737, 0.922)
7	0.651	(0.715, 0.946)		0.713	(0.757, 0.920)
8	0.760	(0.739, 0.956)		0.776	(0.757, 0.938)
9	0.740	(0.754, 0.980)		0.739	(0.731, 0.932)
10	0.667	(0.735, 0.948)		0.705	(0.740, 0.920)
11	0.665	(0.721, 0.947)		0.699	(0.746, 0.922)
12	0.729	(0.649, 0.970)		0.747	(0.692, 0.910)
13	0.773	(0.744, 0.964)		0.759	(0.718, 0.918)
14	0.707	(0.730, 0.956)		0.731	(0.732, 0.923)
16	0.771	(0.736, 0.988)		0.757	(0.722, 0.934)
17	0.696	(0.716, 0.977)		0.759	(0.727, 0.931)
18	0.748	(0.743, 0.977)		0.704	(0.685, 0.917)

Note: AUC values are statistically significant at $p<.05$ if the 95% confidence intervals do not include 0.5. [a] Items 2 and 15 were not administered because they contained the word “suicide.”

Supplemental Table 5.

Percentage of patients with each screening score configuration at baseline, and probability of follow-up suicidal behavior associated with each screening configuration

Screening results[a]	% with suicidal behavior within 30 days	% with suicidal behavior within 90 days
Full sample (i.e., no screening)	0.5	0.5
PHQ-2 positive	1.7	3.2
PHQ-2 positive + PHQ-9 item 9		
PHQ-9 item 9 = 0	0.7	1.0
PHQ-9 item 9 >= 1	4.2	8.4
PHQ-9 item 9 >= 2	5.6	11.1
PHQ-9 item 9 = 3	10.0	16.7
PHQ-2 positive + PHQ-9 >= 1 + SCS item 8		
SCS item 8 = 0	0.0	0.0
SCS item 8 >= 1	4.2	9.1
SCS item 8 >= 2	5.0	10.7
SCS item 8 >= 3	5.9	12.9
SCS item 8 = 4	10.7	17.9
PHQ-2 positive + PHQ-9 >= 1 + SCS item 13		
SCS item 13 = 0	0.0	2.6
SCS item 13 >= 1	5.3	5.2
SCS item 13 >= 2	7.1	6.7
SCS item 13 >= 3	7.4	9.6
SCS item 13 = 4	13.0	17.2
PHQ-2 positive + PHQ-9 >= 1 + SCS item 16		
SCS item 16 = 0	1.9	3.8
SCS item 16 >= 1	5.8	12.6
SCS item 16 >= 2	6.8	13.7
SCS item 16 >= 3	12.0	16.0
SCS item 16 = 4	18.2	27.3

[a] A positive screen on the PHQ-2 corresponds to a total score of 3 or higher, a positive screen on the PHQ-9 corresponds to a score of 1 or higher on item 9, and a positive screen on SCS items 7, 12, and 14 corresponds to a score of 2 or higher on each item.

Supplemental References

1. Rubin DB. Inference and missing data. *Biometrika*. 1976;63:581–90.
2. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley and Sons; 2004.